DIGITALCOMMONS @WAYNESTATE

Journal of Modern Applied Statistical Methods

Volume 8 | Issue 2 Article 23

11-1-2009

Bayesian Analysis of Evidence from Studies of Warfarin v Aspirin for Symptomatic Intracranial Stenosis

Vicki Hertzberg
Emory University, vhertzb@sph.emory.edu

Barney Stern
University of Maryland, bstern@som.umaryland.edu

Karen Johnston
University of Virginia, kj4v@virginia.edu

Journal of Modern Applied Statistical Methods
November 2009, Vol. 8, No. 2, 583-591

1538 – 9472/09/$95.00

Bayesian Analysis of Evidence from Studies of Warfarin v Aspirin for Symptomatic Intracranial Stenosis

Vicki Hertzberg
Emory University

Barney Stern
University of Maryland

Karen Johnston
University of Virginia

Bayesian analyses of symptomatic intracranial stenosis studies were conducted to compare the benefits of long-term therapy with warfarin to aspirin. The synthesis of evidence of effect from previous non-randomized studies in monitoring a randomized clinical trial was of particular interest. Sequential Bayesian learning analysis was conducted and Bayesian hierarchical random effects models were used to incorporate variability between studies. The posterior point estimates for the risk rate ratio (RRR) were similar between analyses, although the interval estimates resulting from the hierarchical analyses are larger than the corresponding Bayesian learning analyses. This demonstrated the difference between these methods in accounting for between-study variability. This study suggests that Bayesian synthesis can be a useful supplement to futility analysis in the process of monitoring randomized clinical trials.

Key words: Bayesian analysis, Bayesian hierarchical model, Bayesian learning, randomized clinical trial, epidemiology, stroke.

Introduction

A responsibility of the committees charged with monitoring randomized clinical trials is to track new evidence from similar studies. However, there are no specific guidelines for the assembly and analysis of such information. Recently the use of Bayesian methods has become accepted in the randomized clinical trials (RCT) community. (Berry, Berry, McKellar & Pearson, 2003). One area in which Bayesian methods are useful is in the synthesis of evidence. (Spiegelhalter, Abrams & Myles, 2004) Thus such methods could provide a data safety committee with useful insights into relevant external information accumulating during the course of a study.

Vicki Hertzberg is an Associate Professor in the School of Public Health, Department of Biostatistics. Email: vhertzb@sph.emory.edu. Barney Stern is a Professor in the School of Medicine, Department of Neurology and Neurosurgery. Email him at: bstern@som.umaryland.edu. Karen Johnston is a Professor in the Department of Neurology. Email: kj4v@Virginia.EDU.

Although Bayesian methods are growing in acceptance in the RCT community, their use in stroke RCTs is still debated (Berry 2005, Donnan, Davis & Ludbrook, 2005; Howard, Coffey & Cutter, 2005; Krams, Lees & Berry, 2005). This study explores the use of two Bayesian techniques for synthesis of evidence. Specifically sequential Bayesian learning and hierarchical Bayesian models are used (Gelman, Carlin, Stern & Rubin, 2004) to examine results from accumulating studies, then illustrate their application to the Warfarin v Aspirin for Symptomatic Intracranial Disease (WASID) trial.

WASID Background

A long-standing secondary stroke prevention strategy for patients with symptomatic intracranial atherostenosis has been warfarin therapy. Warfarin's use was predicated on evidence published in a case series from the Mayo Clinic in the 1950's (Millikan, Siekert & Shick, 1954). This finding was subsequently supported by similarly positive results in observational studies (Marzewski, et al,. 1982; Moufarrij, Little, Furlan, Williams & Marzewski, 1984; Chimowitz, et al., 1995; Thijs & Albers, 2000; Qureshi, et al., 2003)

In 1998, the National Institute of Neurological Diseases and Stroke (NINDS) funded the Warfarin vs Aspirin for Symptomatic Intracranial Disease (WASID) study, the first double-blinded, placebo-controlled randomized clinical trial (RCT) to test the superiority of warfarin (International Normalized Ratio [INR] 2 – 3) over high-dose aspirin (650 mg twice daily) in this patient population. The protocol called for enrollment of 806 patients with angiographically proven symptomatic intracranial disease to determine a combined endpoint of stroke (ischemic and hemorrhagic) and vascular death. The sample size was chosen to give 80% power to detect a difference between event rates of 33% in the aspirin group compared to 22% in the warfarin group over 3 years after, accounting for a 24% rate of discontinuation of study medications and 1% loss to follow-up, which translates to an alternative hazard ratio (HR) of 1.47.

In July, 2003, after 569 patients were enrolled, NINDS, acting upon advice from the WASID Performance and Safety Monitoring Board (PSMB), stopped WASID because subjects randomized to warfarin were at significant increased risk of major non-endpoint adverse events and the potential for a benefit in primary endpoint events that was sufficient to outweigh these adverse events was very low. Indeed, after study closeout, there was no advantage of warfarin versus aspirin (HR = 1.04; 95% CI = 0.73 to 1.48) (Chimowitz, et al., 2005).

Description of Prior Evidence

Existing literature on warfarin treatment for intracranial stenosis was reviewed (Millikan, et al., 1954; Marzewski, et al., 1982; Moufarrij, et al., 1984; Chimowitz, et al., 1995; Thijs, et al., 2000; Qureshi, et al., 2003). Of the six publications, two studies (Marzewski, et al., 1982; Moufarrij, et al., 1984) insufficiently detailed; focus is placed on the remaining four publications in addition to the article describing the WASID trial results (Chimowitz, et al. 2005). (Relevant features of these studies, along with pertinent effect estimates, are summarized in Table 1.)

Study 1: Millikan, et al. (1954) examined Mayo Clinic patients with either intermittent insufficiency of the basilar system or thrombosis within the basilar arterial system. They found that 10/23 (43%) of patients who did not receive anticoagulant therapy died, compared to 3/21 (14%) of patients receiving anticoagulants. The estimated odds ratio (OR) for death comparing aspirin to warfarin (with associated 95% confidence interval [CI]) is 4.62 (2.18, 9.79).

Study 2: Chimowitz, et al. (1995) assessed cases with symptomatic, angiographically confirmed stenosis (≥ 50%) of a major intracranial artery in a retrospective, non-randomized cohort study. Of the 151 patients included in the study, 88 were treated with warfarin and 63 were treated with aspirin. Treatments and dosages were chosen by local physician. Patients were followed by chart review and telephone or personal / next-of-kin interview until first occurrence of a primary endpoint (major vascular event defined as ischemic stroke, myocardial infarction or sudden death), change in therapy (from aspirin to warfarin or vice versa), or last contact or death due to non-vascular cause. Warfarin patients were followed for a median duration of 14.7 months, experiencing 8.4 major vascular events per 100 patient years of follow-up. Aspirin patients were followed for a median duration of 19.3 months, experiencing 18.1 major vascular events per 100 patient years. The estimate of relative risk (RR) of major vascular events in aspirin patients compared to warfarin patients is 2.2 (95% CI, 1.2, 4.4).

Study 3: Thijs and Albers (2000) interviewed 51 patients identified from chart review. All patients had symptomatic intracranial stenosis and had failed antithrombotic therapy. Of these, 32 patients were followed on warfarin and 19 on aspirin. Cox proportional hazards analysis was conducted to estimate the hazard ratio (HR) for cerebral ischemic events (including TIA) after adjusting for age, presence of anterior circulation disease, Caucasian race, and hyperlipidemia. The estimated aspirin to warfarin HR is 4.9 (95% CI, 1.7, 13.9).

Study 4: Qureshi, et al. (2003) retrospectively assessed 102 patients with symptomatic vertebrobasilar stenosis. Cox proportional hazards analysis gave an estimated

Table 1: Data Used in Study Analyses

Study Number & Author(s)	Year	Endpoint	Warfarin: #events/ #observations	Aspirin: #events/ #observations	Aspirin/Warfarin ratio (95% CI)	Log(ratio) and (sd)	Caveat*
(1) Millikan, et al.	1954	Death	3 / 21 patients	10 / 23 patients	4.62 (2.18, 9.79)	1.53 (0.75)	A
(2) Chimowitz, et al.	1995	Stroke, MI, sudden death	26 / 143 patient-year	14 / 166 patient-year	2.17 (1.16, 4.35)	0.63 (0.33)	B
(3) Thijs and Albers	2000	Cerebral ischemic events	Not given	Not given	4.9 (1.7, 13.9)	0.77 (0.33)	C
(4) Qureshi, et al.	2003	Stroke or death	10 / 619 patient-month	8 / 787 patient-month	0.63 (0.25, 1.59)	-0.46 (0.47)	D
(5) Chimowitz, et al.	2005	Ischemic stroke, brain hemorrhage, vascular death	62 / 504.4 patient-year	63 / 541.7 patient-year	1.04 (0.73, 1.48)	0.06 (0.18)	

Caveats:
A: The treatment received by patients not receiving warfarin is unclear as are the inclusion criteria
B: Retrospective study possibly subject to selection bias
C: HRR is adjusted for age, presence of anterior circulation disease, Caucasian race, hyperlipidemia
D: Unpublished result from data supporting paper; Qureshi & Suri, personal communication, December 22, 2005

HR of 55.6 (95% CI, 9.1, 333), comparing stroke free survival for patients receiving either warfarin or aspirin to patients receiving neither after adjustment for sex, race, hypertension, diabetes mellitus, cigarette smoking, hyperlipidemia, and lesion location. Additional data provided by the authors (Table 2) allowed calculation of the aspirin to warfarin HR as 0.63 (95% CI, 0.25, 1.59).

Study 5: Chimowitz, et al. (2005) was the only RCT comparing warfarin to aspirin in patients with this disease. 569 patients were followed for an average of 1.8 years. The aspirin to warfarin HR is 1.04 (95% CI, 0.74, 1.49).

Table 2: Endpoints (Stroke or Death) in Qureshi, et al. (2003)*

	Warfarin (n=46)	Aspirin (n =40)
Number of Patients	46	40
Stroke or Death	10	8
Person-Months to Endpoint	619	787

*Qureshi & Suri, personal communication, December 22, 2005

Methodology

Bayesian Learning

Because the results of these studies were accumulated over 50 years, a Bayesian learning approach was first used in which the posterior distribution derived from the analysis of the oldest result was used as the prior distribution in order to derive the posterior distribution with the next study. The goal was to estimate the posterior distribution of θ, the unknown mean of the distribution of log(RRR) from its prior and the preceding study results with the posterior distribution derived from study *i-1* serving as the prior distribution for study *i* for $i = 2,\ldots 5$. This is expressed as follows:

Let $Y_i = \log(RRR_i)$.

Assuming that Y_i is a realization from a random distribution depending on θ, the Bayes theorem gives

$$f(\theta|Y_1) \propto f(Y_1|\,\theta) \times f(\theta), \qquad (1)$$

$$f(\theta|Y_2) \propto f(Y_2|\,\theta) \times f(\theta|Y_1). \qquad (2)$$

In general $f(\theta|Y_i) \propto f(Y_i|\,\theta) \times f(\theta|Y_{i-1})$, where $f(\theta)$ is the baseline prior distribution for θ and $i > 1$, assuming that log(RRR) is normally distributed, using the normal distribution for the likelihood and its conjugate, the normal distribution, as the prior for θ.

Hierarchical Random Effects Models

A simultaneous analysis in a hierarchical random effects model was also considered, specifically each has an estimate Y_i of a treatment effect θ_i, such that:

$$Y_i \sim f(y_i \mid \theta_i). \qquad (3)$$

These treatment effects are treated as realizations of random variables from the same population, that is,

$$\theta_i \sim f(\theta_i \mid \theta_\mu), \qquad (4)$$

with θ_μ having its own prior distribution $f(\theta_\mu)$.

Because all of the studies present an estimate of risk which, after transformation, has a normal distribution, a normal distribution was used for functional forms of likelihood and prior distribution functions. For some studies the results may also be viewed as events per person-years of observation per group. In this case Poisson hierarchical models can be used as follows:

> For group j in study i let the number of events, $E_{ij} \sim \text{Poiss}\,(n_{ij} \times \exp(\varphi x_j + \varepsilon_{ij}))$, where x_j is an indicator variable denoting aspirin group membership, $\varepsilon_{ij} \sim N(0, \sigma^2_\varepsilon)$ and $\varphi \sim N(0, \sigma^2_\varphi)$. Here n_{ij}, the number of person years at risk, is an offset term and φ is the population value for log(RRR).

For each of these analyses a posterior mean was generated with 95% Bayes interval and posterior median with 50% Bayes interval or inter-quartile range.

For each analysis three different baseline priors were used as follows:

1) $\theta \sim N(0, 10)$ (a weakly non-informative prior; warfarin has no effect);
2) $\theta \sim N(0, 0.5)$ (a skeptical prior; warfarin has no effect);
3) $\theta \sim N(0.5, 10)$ (an enthusiastic prior; warfarin reduces risk by 40%).

Additional sensitivity analyses included only studies 2, 4 and 5.

As the Bayesian learning analysis proceeded, graphs of the posterior, likelihood and prior functions were inspected at each step. Thus, the relative influence that the likelihood and prior exerted in determination of the resulting posterior was able to be determined.

Numerical Methods

In addition to the distributions for the parameters of interest, non-informative prior distributions were placed on any nuisance parameters (e.g., σ^2_ε in the hierarchical Poisson model) then integrated over these parameters in the posterior distribution. For estimation, Gibbs sampling (Casella & George, 1992) was used as performed in the WinBUGS software (Spiegelhalter, Thomas, Best, Gilks & Lunn, 2003). Convergence was monitored using the scale reduction factor (SRF) (Gelman, et al., 2004). For each model analyzed, 3 chains were run with 1,000 iterations each (discarding the first 500 in each chain). For analyses which resulted in SRF > 1.1 the number of iterations was increased in each chain by a factor of 10 the

program run again until the SRF ≤ 1.1. Note that such increases were only necessary for analysis of the hierarchical random effects Poisson models.

Results

Bayesian Learning Analyses

The results of the sequential Bayesian learning analysis with log(RRR) as a normal variate using studies 1-4 are shown in Table 3. Note that the Bayesian results that were available at the time that WASID began (studies 1 and 2) were mixed in their support for an effect of warfarin as hypothesized for the WASID clinical trial, i.e., RRR = .33 / .22 = 1.5, versus the null hypothesis RRR = 1.

Specifically, although the 95% Bayes intervals based on the initial informative prior or the initial enthusiastic prior include 1.5 but exclude 1, the interval based on the initial skeptical prior includes both values. With the subsequent addition of study 3's results the evidence favoring warfarin grew stronger. The 95% Bayes intervals stemming from both initial skeptical and initial enthusiastic priors now include 1.5 but exclude 1. Moreover the interval stemming from the initial non-informative prior excludes both 1 and 1.5 to the left. Addition of study 4 has little effect on point and interval estimates. Further point estimates stemming from the non-informative prior tend to be much higher than corresponding point estimates stemming from skeptical and enthusiastic priors at each point. This disparity is due to the difference in variance between the non-informative prior versus skeptical and enthusiastic priors.

A hypothetical future study of warfarin and aspirin would incorporate the results of study 5. With this addition, note that interval estimates stemming from initial non-informative and skeptical priors now include 1. Indeed, the Bayes interval from the initial skeptical prior now excludes 1.5 to the right. The Bayes interval stemming from the enthusiastic prior excludes 1 but covers 1.5 (the alternative hypothesis for WASID) and 2.

Sensitivity analyses including only studies 2, 4 and 5 result in posterior point and interval estimates that are not much different after adding study 4 into the analysis, especially with the skeptical and enthusiastic priors (Table 4). The results after introduction of only study 2 are like the results after inclusion of both studies 1 and 2, suggesting that the optimistic estimates from study 1 do not contribute substantially to the overall conclusion. Additional sensitivity analysis including study 4 produced posterior point and interval estimates that were virtually identical suggesting that study 4 does not have a substantial impact on the analysis.

Table 3: Posterior Point and Interval Estimates for θ = RRR from Bayesian Learning Analysis Using All Studies

Study Added	Interval Type	Non-Informative Prior	Skeptical Prior	Enthusiastic Prior
1	Mean θ (95% Bayes interval)	4.48 (1.14, 17.64)	1.11 (0.75, 1.63)	1.82 (1.23, 2.69)
	Median θ (50% Bayes interval)	4.48 (2.72, 7.39)	1.22 (1.0, 1.35)	1.82 (1.49, 2.23)
2	Mean θ (95% Bayes interval)	2.46 (1.36, 4.44)	1.35 (0.91, 1.99)	1.82 (1.23, 2.69)
	Median θ (50% Bayes interval)	2.46 (2.01, 3.00)	1.35 (1.22, 1.49)	1.82 (1.65, 2.01)
3	Mean θ (95% Bayes interval)	3.00 (1.67, 5.42)	1.65 (1.12, 2.44)	2.01 (1.36, 2.97)
	Median θ (50% Bayes interval)	3.00 (2.46, 3.67)	1.65 (1.49, 1.82)	2.01 (1.82, 2.46)
4	Mean θ (95% Bayes interval)	2.72 (1.65, 4.95)	1.65 (1.11, 2.46)	2.01 (1.35, 3.00)
	Median θ (50% Bayes interval)	2.72 (2.23, 3.32)	1.65 (1.49, 1.82)	2.01 (1.82, 2.23)
5	Mean θ (95% Bayes interval)	1.35 (1.00, 2.01)	1.35 (1.00, 1.43)	1.49 (1.11, 2.01)
	Median θ (50% Bayes interval)	1.49 (1.22, 1.65)	1.35 (1.22, 1.49)	1.49 (1.35, 1.65)

Simultaneous Analysis of RRR Using Hierarchical Random Effects Models with the Normal Distribution

The simultaneous analysis of these studies was examined in the normal model for log(RRR). Posterior point and interval estimates for the analyses of various subsets of studies are shown in Table 5. The results are very similar to the results of the comparable Bayesian learning analysis, although with wider intervals, indicating different consequences of the ways these methods address variability between studies. Specifically the Bayesian learning analysis provides for a posterior variance estimate at each step, but this estimate can drift between steps. In comparison, the simultaneous nature of the hierarchical models requires adjustment over all studies at once.

Simultaneous Analysis of RRR Using Hierarchical Random Effects with the Poisson Distribution

The HR from a Poisson model that compares events per person year between thegroups was examined, with interest in the ratio between the two Poisson parameters, which are an estimate of RRR. Since not all studies were sufficiently detailed in their report of rates, these analyses are limited. Nevertheless the extent of knowledge for 3 of the existing studies and subsets was examined.

The results of these analyses are shown in Table 6. The first analysis, using only study 2, represents a simple Bayesian analysis using a Poisson distribution. Note that the value of 1.5 is included in the 95% Bayes interval estimates, while the null value 1.0 is excluded by the analysis using the non-informative and enthusiastic priors. The other two analyses used a hierarchical random effects Poisson model to adjust for differences between studies. In these analyses using studies 2 and 4 or using studies 2, 4, and 5, the 95% Bayes interval estimates include both 1 and 1.5.

Table 4: Posterior Point and Interval Estimates for θ = RRR from Bayesian Learning Analysis Using Restricted Set of Studies

Study Added	Interval Type	Non-Informative Prior	Skeptical Prior	Enthusiastic Prior
2	Mean θ (95% Bayes interval)	2.23 (1.23, 4.01)	1.35 (0.91, 1.99)	1.82 (1.23, 2.69)
	Median θ (50% Bayes interval)	2.23 (1.82, 2.72)	1.35 (1.22, 1.65)	1.82 (1.65, 2.01)
4	Mean θ (95% Bayes interval)	2.01 (1.22, 3.67)	1.35 (0.90, 2.01)	1.82 (1.22, 2.72)
	Median θ (50% Bayes interval)	2.23 (1.82, 2.46)	1.35 (1.22, 1.49)	1.82 (1.65, 2.01)
5	Mean θ (95% Bayes interval)	1.35 (0.90, 1.82)	1.22 (0.90, 1.65)	1.35 (1.11, 1.82)
	Median θ (50% Bayes interval)	1.35 (1.11, 1.49)	1.22 (1.11, 1.35)	1.35 (1.22, 1.49)

Table 5: Posterior Point and Interval estimates for θ = RRR Using Hierarchical Random Effects Model with Normal Distribution

Studies Included	Interval Type	Non-Informative Prior	Skeptical Prior	Enthusiastic Prior
1, 2, 3, 4	Mean θ (95% Bayes interval)	2.72 (0.27, 14.88)	1.11 (0.67, 1.82)	1.82 (1.22, 3.00)
	Median θ (50% Bayes interval)	3.00 (2.01, 4.06)	1.11 (1.00, 1.35)	1.82 (1.65, 2.23)
1, 2, 3, 4, 5	Mean θ (95% Bayes interval)	2.01 (0.55, 6.69)	1.22 (0.74, 1.82)	1.82 (1.22, 2.72)
	Median θ (50% Bayes interval)	2.01 (1.49, 2.72)	1.22 (1.00, 1.35)	1.82 (1.49, 2.01)
2, 4	Mean θ (95% Bayes interval)	1.49 (0, 4.9×10^5)	1.00 (0.67, 1.65)	1.65 (1.00, 2.72)
	Median θ (50% Bayes interval)	1.82 (0.27, 7.39)	1.00 (0.90, 1.22)	1.65 (1.35, 2.01)
2, 4, 5	Mean θ (95% Bayes interval)	1.22 (0.07, 20.1)	1.11 (0.67, 1.65)	1.49 (1.11, 2.46)
	Median θ (50% Bayes interval)	1.35 (0.90, 2.01)	1.11 (1.00, 1.22)	1.49 (1.35, 1.82)

Conclusion

Reconciliation of the results of WASID with previous reports of a strong effect of warfarin is difficult. Many would advocate that the biases of the prior observational studies should discount those results in favor of the unbiased result of the RCT. Certainly the use of randomization, blinding, standardization of patient management protocols, and central endpoint adjudication ensure bias-free estimate of treatment effect from the RCT. However, RCTs are not without other sources of bias stemming from the selection of participating physicians and clinics as well as the enrollment of consenting patients. Thus, a growing community of investigators (Berry, et al., 2003; Brophy & Lawrence, 1995; Diamond & Kaul, 2004) advocates the use of Bayesian statistical methods to interpret results of clinical trials as well as to synthesize evidence from a set of studies about the effect of treatment(s). Bayesian statistical methods have recently gained notice in the arena of stroke clinical trials (Berry, 2005; Donnan, et al., 2005; Howard, et al., 2005; Krams, et al., 2005).

Although taken as a single trial the WASID results would seem to extinguish the utility of warfarin as a secondary prevention strategy for patients with symptomatic warfarin's demise (Koroshetz, 2005). In this presentation we explore application of Bayesian methods to interpret the WASID results in light of the overall accumulation of evidence regarding the effect of warfarin and consider what insights the Bayesian analyses might have indicated along the way?

At the time of the WASID proposal submission, the accumulated evidence taken from the Bayesian learning perspective fit neatly with the standard of equipoise necessary to justify NIH funding. Specifically, those coming to the debate with no or vague prior beliefs (i.e., the non-informative prior) as well as those favoring warfarin (i.e., the enthusiast) could justify RRR = 1.5 and exclude RRR = 1. On the other hand, those coming to the problem favoring no difference (i.e., the skeptic) could justify both values for RRR. With the hierarchical analyses the alignments of skeptics and enthusiasts remain the same, while those with vague beliefs now align with the skeptics.

In July 2003, when the study was terminated for safety reasons, the results of the Bayesian learning analyses all excluded RRR = 1 from interval estimates, regardless of prior beliefs. When the analysis is restricted to studies meeting perceived quality criteria, the initial state of equipoise described above remained.

Moreover, the hierarchical analyses limited to published results as of July 2003 would be no different than before. However, the inclusion of the rates from study 4, if they had been published at that time, leads to hierarchical model results that lend support for both RRR=1.5 or RRR=1 regardless of prior belief.

The lack of strict correspondence between conclusions from Bayesian learning with those from Bayesian hierarchical random effects models results from differences between methods in incorporating between-study variability. The studies do have differences in design (sample size, endpoint definitions and inclusion criteria) warranting allowances in the modeling process. Although none of studies 1-4 were randomized clinical trials, hierarchical models can be extended to adjust for different classes (such as RCTs versus non-randomized studies) when 2 or more studies of each class are present. Unfortunately only one RCT was available to include.

It is particularly interesting to note the change in conclusions wrought by the unpublished, negative result of Study 4. This finding reinforces the importance of finding all results, even negative ones, in compiling evidence.

The ability to generate interval estimates and use differing priors deepens understanding of the current evidence in light of previous studies. These results point to the utility of Bayesian analyses of prior studies as an additional tool for monitoring clinical trials. The concordance of frequentist and Bayesian efficacy analyses would provide robust confirmation of the appropriateness of a futility analysis when decisions regarding the continuation or stopping of a clinical trial are made.

Acknowledgements

This study was funded by a research grant (1R01 NS33643) from the US Public Health Service National Institute of Neurological Disorders and Stroke (NINDS). We thank Marc Chimowitz, Mike Lynn, Scott Janis, and Bill Powers for their careful review and comments.

References

Berry, D. (2005). Clinical trials: Is the Bayesian approach ready for prime time? Yes! *Stroke*, *36*, 1621-1623.

Table 6: Posterior Point and Interval estimates for φ = RRR Using Hierarchical Random Effects Model with Poisson Distribution

Studies Included	Interval Type	Non-Informative Prior	Skeptical Prior	Enthusiastic Prior
2*	Mean φ (95% Bayes interval)	2.23 (1.22, 4.48)	1.65 (1.0, 3.0)	2.01 (1.22, 3.32)
	Median φ (50% Bayes interval)	2.23 (1.82, 2.72)	1.82 (1.49, 2.01)	2.01 (1.65, 2.46)
2, 4	Mean φ (95% Bayes interval)	0.41 (0, 54200)	1.0 (0.41, 2.72)	1.65 (0.67, 4.48)
	Median φ (50% Bayes interval)	0.41 (0.01, 22.20)	1.0 (0.74. 1.35)	1.65 (1.22, 2.23)
2, 4, 5	Mean φ (95% Bayes interval)	1.11 (0, 24300)	1.0 (0.41, 2.72)	1.65 (0.55, 4.48)
	Median φ (50% Bayes interval)	1.11 (0.03, 54.6)	1.0 (0.74, 1.35)	1.65 (1.11, 2.23)

*Uses simple Poisson model for two groups

Berry, D., Berry, S., McKellar, J., & Pearson, T. (2003). Comparison of the dose-response relationships of 2 lipid-lowering agents: A Bayesian meta-analysis. *American Heart Journal*, *145*, 1036-1045.

Brophy, J., & Lawrence, J. (1995). Placing trials in context using Bayesian analysis: Gusto revisited by reverend Bayes. *JAMA*, *273*, 871-875.

Casella, G., & George, E. (1992). Explaining the Gibbs sampler. *American Statistician*, *46*, 167-174.

Chimowitz, M., et al. (1995). The warfarin-aspirin symptomatic intracranial disease study. *Neurology*, *45*, 1488-1493.

Chimowitz, M., et al. (2005). Comparison of warfarin and aspirin for symptomatic intracranial arterial stenosis. *New England Journal of Medicine*, *352*, 1305-1316.

Diamond, G., & Kaul, S. (2004). Prior convictions: Bayesian approaches to the analysis and interpretation of clinical megatrials. *Journal of the American College of Cardiology*, *43*, 1929-1939.

Donnan, G., Davis, S., & Ludbrook, J. (2005). The Bayesian principle: Can we adapt? *Stroke*, *36*, 1623-1624.

Gelman, A., Carlin, J., Stern, H., & Rubin, D. (2004). *Bayesian data analysis* (*2nd Ed.*). NY: Chapman & Hall.

Howard, G., Coffey, C., & Cutter, G. (2005). Is Bayesian analysis ready for use in phase iii randomized clinical trials? Beware the sound of the sirens. *Stroke*, *36*, 1622-1623.

Investigators, T.W.-A.S.I.D.W.T. (2003). Design, progress, and challenges of a double-blind trial of warfarin versus aspirin for symptomatic intracranial arterial stenosis. *Neuroepidemiology*, *22*, 106-117.

Koroshetz, W. (2005). Warfarin, aspirin, and intracranial vascular disease. *New England Journal of Medicine*, *352*, 1368-1370.

Krams, M., Lees, K., & Berry, D. (2005). The past is the future: Innovative designs in acute stroke therapy trials. *Stroke*, *36*, 1341-1347.

Marzewski, D., et al. (1982). Intracranial internal carotid artery stenosis: Longterm prognosis, *Stroke*, *13*, 821-824.

Millikan, C., Siekert, R., & Shick, R. (1954). *Studies in cerrebrovascular disease. Iii. The use of anticoagulant drugs in the treatment of insufficiency or thrombosis within the basilar artery system*. Staff Meetings of the Mayo Clinic, *30*, 116-126.

Moufarrij, N., Little, J., Furlan, A., Williams, G., & Marzewski, D. (1984). Vertebral artery stenosis: Long-term follow-up. *Stroke*, *15*, 260-263.

Qureshi, A., et al. (2003). Stroke-free survival and its determinants in patients with symptomatic vertebrobasilar stenosis: A multicenter study. *Neurosurgery*, *52*, 1033-1040.

Spiegelhalter, D., Abrams, K., & Myles, J. (2004). *Bayesian approaches to clinical trials and health-care evaluation*. Chichester, England: John Wiley & Sons.

Spiegelhalter, D., Thomas, A., Best, N., Gilks, W., & Lunn, D. (2003). Bugs: Bayesian *Inference using Gibbs sampling, Technical, MRC Biostatistics Unit*.

Thijs, V., & Albers, G. (2000). Symptomatic intracranial atherosclerosis: outcome of patients who fail antithrombotic therapy. *Neurology*, *55*, 490-498.